

# Mass models of the Milky Way

Paul J. McMillan<sup>\*</sup>

*Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, Oxford, OX1 3NP, UK*

12 January 2013

## ABSTRACT

We present a simple method for fitting parametrized mass models of the Milky Way to observational constraints. We take a Bayesian approach which allows us to take into account input from photometric and kinematic data, and expectations from theoretical modelling. This provides us with a best-fitting model, which is a suitable starting point for dynamical modelling. We also determine a probability density function on the properties of the model, which demonstrates that the mass distribution of the Galaxy remains very uncertain. For our choices of parametrization and constraints, we find disc scale lengths of  $3.00 \pm 0.22$  kpc and  $3.29 \pm 0.56$  kpc for the thin and thick discs respectively; a Solar radius of  $8.29 \pm 0.16$  kpc and a circular speed at the Sun of  $239 \pm 5$  km s<sup>−1</sup>; a total stellar mass of  $6.43 \pm 0.63 \times 10^{10} M_{\odot}$ ; a virial mass of  $1.26 \pm 0.24 \times 10^{12} M_{\odot}$  and a local dark matter density of  $0.40 \pm 0.04$  GeV cm<sup>−3</sup>. We find some correlations between the best-fitting parameters of our models (for example, between the disk scale lengths and the Solar radius), which we discuss. The chosen disc scale-heights are shown to have little effect on the key properties of the model.

**Key words:** Galaxy: fundamental parameters – methods: statistical – Galaxy: kinematics and dynamics

## 1 INTRODUCTION

A great deal is still unknown about the distribution of mass in the various components of the Milky Way. The major discoveries in Galactic astronomy over the past decade have almost all been related to components which comprise a small fraction of the total mass of the Milky Way, most of which either are or were dwarf galaxies (for example, the many objects observed in the “Field of Streams”, Belokurov et al. 2006). The structure of the dominant components – the disc(s) and cold dark matter (CDM) halo – remains rather uncertain.

An important element of understanding and constraining the structure of the major components of the Galaxy is creating Galaxy models which can be compared to observational data. It is important to draw a distinction between three types of Galaxy models: mass, kinematic and dynamical models. Mass models are the simplest of these, and only attempt to describe the density distribution of the various Galaxy components, and thus the Galactic potential (e.g. Klypin, Zhao, & Somerville 2002; Dehnen & Binney 1998, henceforth DB98). Kinematic models, such as those produced by GALAXIA (Sharma et al. 2011), specify the density and velocity distributions of the luminous components of the Galaxy, but do not consider the question of whether

these are consistent with a steady state in any Galactic potential. Dynamical models (e.g. Widrow, Pym, & Dubinski 2008) describe systems which are in a steady state in a given potential, because their distribution functions depend only on the integrals of motion. The Besançon Galaxy model (Robin et al. 2003) is primarily a kinematic model with a dynamical element used to determine the vertical structure of the disc.

It is clear that moving beyond simple kinematic models to full dynamical ones is an essential step in fully understanding our Galaxy. The majority of the mass of the Galaxy is expected to lie in the CDM halo, which is only observable through its gravitational effect on luminous components of the Galaxy, so purely kinematic models cannot provide any insight into its structure.

The first step towards a dynamical model is to produce a mass model that is consistent with available constraints. An influential mass model was that of Schmidt (1956), and, as observational data and understanding of galaxy structure improved, updated versions have been produced by other authors, notably Caldwell & Ostriker (1981) and DB98. Our intention in this study is to follow these authors in producing a mass model that is consistent with up-to-date observational data and theoretical understanding, and to provide a simple framework for producing these models into which future data can be placed as they become available.

The major difficulty in producing a model of this kind is drawing together data from numerous different studies of

<sup>\*</sup> E-mail: p.mcmillan1@physics.ox.ac.uk

the various components which make up the Milky Way in a way that is as consistent as possible. Such studies often make different underlying assumptions, and sometimes come to seemingly mutually contradictory conclusions. In principle the correct approach is to return to the raw data that each study was based upon and to synthesise them into one coherent picture. Even if this is possible in practice, it is certainly an immense undertaking, and one we do not attempt here. Instead we follow the approach of previous authors by accepting various constraints on the parameters of our models as stated by other studies, without returning to the raw data. In addition we use well understood kinematic data sets, and – for the CDM halo, about which observational data is limited – we use our best current theoretical understanding. Our approach is best thought of as Bayesian with direct constraints on the model parameters (from photometric data or theoretical insight) being our Bayesian “prior”, and kinematic data used to find the likelihood.

In this paper we present a simple method for determining both a best-fitting parametrized mass model of the Galaxy, and the full probability density function of the parameters of the model. This paper is associated with a series of papers in which we construct dynamical models to fit observational data (Binney 2010; Binney & McMillan 2011, McMillan et al. in preparation). Eventually these dynamical models will themselves be used to constrain the Galactic potential.

Like DB98 we restrict ourselves to axisymmetric models. It is clear that the Galaxy is not actually axisymmetric, especially the inner Galaxy (see Section 2.1), but the disc is close to axisymmetric (Jurić et al. 2008), and (as noted by DB98) axisymmetric models successfully account for observations in the 21-cm line of hydrogen at Galactic longitudes  $l \gtrsim 30$ , corresponding to  $R \gtrsim 4$  kpc.

To find the gravitational potential associated with a given mass model we use the publicly available code GALPOT, which is described by DB98 section 2.3.

In Section 2 we describe the different components that make up the bulk of the mass of the Galaxy, and how our model represents them. We also explain our Bayesian priors on the parameters that describe our model. In Section 3 we give the kinematic data we use to constrain our model. All the constraints applied to the model are summarised in Table 1. In Section 4, we describe the method used to fit the model, and in Section 5 we give our best-fitting model and the rest of our results. In Section 6 we compare our results to other data sets.

Unless otherwise stated, any measurement or derived quantity we use to constrain our model is assumed to have Gaussian uncertainties.

## 2 COMPONENTS OF THE MILKY WAY

### 2.1 The bulge

The Galactic bulge has been shown in many studies to be a near-prolate, triaxial rotating bar with its long axis in the plane of the Galaxy (e.g. Binney, Gerhard, & Spergel 1997). However, our interest in this study is in producing an axisymmetric model, so we are forced to make crude approximations in our modelling of this component. We must

therefore accept that it is unwise to compare our model to measurements taken from the inner few kpc of the Galaxy, as it can be expected to do a poor job of reproducing them.

Insight into the structure of the bulge can be gained from photometric studies, however one must be careful when doing so as there can be a major contribution to the stellar density in the inner few kpc from the disc component. The model used for the disc in these studies will therefore have a significant effect on the properties determined for the bulge. This goes some way towards explaining why the mass of the bulge as determined by Picaud & Robin (2004),  $2.4 \pm 0.6 \times 10^{10} M_\odot$ , using a photometric model with a “hole” in the centre of the disc, is so much larger than that determined by studies using kinematic data (e.g. DB98; Widrow, Pym, & Dubinski 2008; Bissantz & Gerhard 2002) which do not include a disc “hole” in their models.

Our density profile is based on the parametric model which Bissantz & Gerhard (2002) fit to dereddened L-band COBE/DIRBE data (Spergel, Malhotra, & Blitz 1996), and the mass-to-light ratio that Bissantz & Gerhard determine from a comparison between gas dynamics in models and those observed in the inner Galaxy. This model is found on the assumption that the disc component has no central hole. This also assumes that the mass-to-light ratio is spatially constant, which allows us to convert a photometric model directly into a mass model for this component.

The Bissantz & Gerhard model is not axisymmetric, so we make an axisymmetrised approximation which has the density profile

$$\rho_b = \frac{\rho_{b,0}}{(1 + r'/r_0)^\alpha} \exp \left[ - (r'/r_{\text{cut}})^2 \right], \quad (1)$$

where, in cylindrical coordinates,

$$r' = \sqrt{R^2 + (z/q)^2} \quad (2)$$

with  $\alpha = 1.8$ ,  $r_0 = 0.075$  kpc,  $r_{\text{cut}} = 2.1$  kpc, and axis ratio  $q = 0.5$ . The Bissantz & Gerhard mass-to-light ratio has a quoted uncertainty  $\pm 5\%$ , but given the extent to which we have altered this model it is prudent to recognize the need to introduce further uncertainty in our model fitting. We therefore assume that the bulge mass  $M_b = 8.9 \times 10^9 M_\odot$ , with uncertainty  $\pm 10\%$ . For this density profile, this corresponds to scale density  $\rho_{b,0} = 9.93 \times 10^{10} M_\odot \text{ kpc}^{-3} \pm 10\%$ .

### 2.2 The disc

The Milky Way’s disc is usually considered to have two major components: a thin disc and a thick disc (e.g. Gilmore & Reid 1983). These are generally modelled as exponential in the sense that

$$\rho_d(R, z) = \frac{\Sigma_{d,0}}{2z_d} \exp \left( -\frac{|z|}{z_d} - \frac{R}{R_d} \right), \quad (3)$$

with scale height  $z_d$ , scale length  $R_d$  and central surface density  $\Sigma_{d,0}$ . The total mass of a disc like this is  $M_d = 2\pi \Sigma_{d,0} R_d^2$ .

As mentioned in Section 2.1, some studies of the inner Galaxy prefer models with a central “hole” in the disc (e.g. Picaud & Robin 2004). We do not consider such models here. The kinematic data we consider (Section 3) is all related to parts of the Galaxy which lie outside any central disc hole, so, in this study, a model with a central hole

should simply redistribute mass in the inner few kpc from the disc to the bulge (note that our prior on the bulge would have to be replaced, as it is taken from a study which used a model of the disc with no central hole).

The Jurić et al. (2008) analysis of data from the Sloan Digital Sky Survey (SDSS: Abazajian et al. 2009) showed that the approximation to exponential profiles is a sensible one for the Milky Way, and produced estimates based on photometry for the scale lengths, scale heights and relative densities of the two discs.

The scale-heights of the discs are not at all well constrained by the kinematic data we use in this study, so initially we accept without question the Jurić et al. (2008) best-fitting values,  $z_{d,thin} = 300$  pc and  $z_{d,thick} = 900$  pc. In Section 5.1 we explore the (relatively small) effects of changing the assumed disc scale-heights.

The Jurić et al. scale-lengths for the thin and thick discs are 2.6 and 3.6 kpc respectively, with a quoted uncertainty of 20% in each case. The local density normalisation  $f_{d,\odot} = \rho_{thick}(R_\odot, z_\odot)/\rho_{thin}(R_\odot, z_\odot)$  is quoted as 0.12, with uncertainty 10%. We approximate these uncertainties as Gaussian and uncorrelated, and take these as prior probability distributions on  $R_{d,thin}$ ,  $R_{d,thick}$  and  $f_{d,\odot}$ . Again, we assume a constant mass-to-light ratio which allows us to convert these photometric constraints directly into constraints on the mass density.

### 2.3 The dark-matter halo

For obvious reasons, we cannot use photometric data to constrain the shape of the dark matter distribution, so instead we look to cosmological simulations for insight.

In cosmological simulations that only include dark matter, halo density profiles are well fit by a universal profile, known as the NFW profile (Navarro, Frenk, & White 1996)

$$\rho_h = \frac{\rho_{h,0}}{x(1+x)^2}, \quad (4)$$

where  $x = r/r_h$ , with  $r_h$  the scale radius. It is clear that the baryons within CDM haloes will have an impact on the halo profile, but the nature of this impact remains uncertain because it depends on the complex physics of baryons. Cosmological simulations suggest that the condensation of baryons to the centre of galactic haloes will cause the density profile in the inner parts of the halo to be steeper than  $\rho \propto r^{-1}$ , as it is in the NFW case (Duffy et al. 2010). Observations of low surface brightness galaxies, however, appear to indicate that they lie in dark-matter haloes with constant density cores. This difference between observation and simulations (with or without baryonic physics) is known as the “cusp-core problem” (for an overview see de Blok 2010).

Observations of low surface brightness galaxies and dwarf galaxies can be used to provide more reliable information about the inner dark matter profile than galaxies like the Milky Way because they are dark matter dominated right to their centre, so the mass contribution of the baryons can be modelled with little uncertainty compared to the total mass. The Milky Way is not dark matter dominated and it is very hard to determine slope of the dark matter density profile in the inner Galaxy because the baryonic density is dominant in the inner Galaxy, and still uncertain.

Haloes in dark-matter-only cosmological simulations tend to be significantly prolate, but with a great deal of variation in axis ratios (e.g. Allgood et al. 2006). It is recognised that, again, baryonic physics will play an important role – condensation of baryons to the centres of haloes is expected to make them rounder than dark-matter-only simulations would suggest (Debattista et al. 2008). The shape of the Milky Way’s halo is still very much the subject of debate, with different efforts to fit models of the Sagittarius dwarf’s orbit favouring conflicting halo shapes (see e.g. Law, Majewski, & Johnston 2009).

In this study we do not attempt to address the question of the shape of the Milky Way halo, or the impact of baryonic physics on the CDM density profile. We take the simple model of a spherically symmetric NFW halo (eq. 4). As better understanding of the effect of baryonic physics on CDM haloes or constraints from observations become available, these simple assumptions will have to be revisited.

It is common to describe CDM haloes in terms of their virial mass  $M_{vir}$  – which is the mass contained within the virial radius  $r_{vir}$  – and the concentration  $c = r_{-2}/r_{vir}$ , where  $r_{-2}$  is the radius at which the logarithmic slope of the density profile  $d \log \rho / d \log r = -2$  (for an NFW profile,  $r_{-2} = r_h$ ). The virial radius  $r_{vir}$  is defined as the radius of a sphere centred on the halo centre which has an average density of  $\Delta$  times the critical density, but the definition of  $\Delta$  varies between authors. The definitions that are relevant to this paper are  $\Delta = 200$  (when using this definition we will use the notation  $M_v$ ,  $r_v$  and  $c_v$  for the virial mass, radius, and the concentration), and a value of  $\Delta$  which varies with redshift, derived from spherical top-hat collapse models (e.g. Gunn & Gott 1972), with  $\Delta \approx 94$  today (for which we use the notation  $M_{v'}$ ,  $r_{v'}$ ,  $c_{v'}$ ).

Guo et al. (2010) compared the distribution of halo virial masses found in the Millennium and Millennium-II simulations (Springel et al. 2005; Boylan-Kolchin et al. 2009) to the distribution of galaxy stellar masses found by Li & White (2009) using Sloan Digital Sky Survey data (Abazajian et al. 2009). Guo et al. (2010) found that the ratio of total stellar mass  $M_*$  to halo mass  $M_v$  was well fit by

$$M_* = M_v \times A \left[ \left( \frac{M_v}{M_0} \right)^{-\alpha} + \left( \frac{M_v}{M_0} \right)^\beta \right]^{-\gamma}, \quad (5)$$

with an intrinsic scatter of order 0.2 in  $\log_{10} M_*$ , where  $A = 0.129$ ,  $M_0 = 10^{11.4} M_\odot$ ,  $\alpha = 0.926$ ,  $\beta = 0.261$  and  $\gamma = 2.440$ .

To gain insight into the likely value of the concentration parameter  $c_{v'}$  we consider the study of Boylan-Kolchin et al. (2010). This examined haloes taken from the Millennium-II simulations at redshift zero, in the mass range  $10^{11.5} \leq M_v/[h^{-1} M_\odot] \leq 10^{12.5}$  – a mass range that the Milky Way’s halo is likely to lie in – and determined that the probability distribution of the concentration was well fit by a Gaussian distribution in  $\ln c_{v'}$ , with

$$\ln c_{v'} = 2.256 \pm 0.272. \quad (6)$$

Cosmological simulations predict that typical halo concentration varies with halo mass, but typically only weakly (e.g.  $c_v \propto M_v^{-0.1}$ , Neto et al. 2007) when compared to the intrinsic scatter in  $c$  at a given virial mass, so we accept the relationship given by Boylan-Kolchin et al. (2010) as stated,

independent of  $M_{v'}$ . Baryonic physics is likely to have an effect on the concentration, much as it does on the inner density profile (Duffy et al. 2010), but we neglect that in this study.

The Galaxy’s dark-matter halo is far more massive than its stellar halo, which has a mass  $\sim 4 \times 10^8 M_\odot$  (Bell et al. 2008). We therefore treat the stellar halo as a negligible fraction of the Galactic halo, and do not consider it further.

### 3 KINEMATIC DATA

#### 3.1 The Solar position and velocity

If we do not know the position and velocity of the Sun it is extremely difficult to interpret any other kinematic observations of the Milky Way. Unfortunately there is still significant uncertainty on the key parameters, namely the distance from the Sun to the Galactic Centre  $R_0$ , the rotational speed of the local standard of rest (LSR)  $v_0$ , and the velocity of the Sun with respect to the LSR  $\mathbf{v}_\odot$  (see e.g. McMillan & Binney 2010). In this work we take  $R_0$  from the study of stellar orbits around the supermassive black hole at the Galactic Centre, Sgr A\*, by Gillessen et al. (2009)

$$R_0 = 8.33 \pm 0.35 \text{ kpc}, \quad (7)$$

though we could equally have taken the value  $8.4 \pm 0.4 \text{ kpc}$  from Ghez et al. (2008), found from similar data.

A series of recent papers (Binney 2010; McMillan & Binney 2010; Schönrich, Binney, & Dehnen 2010) have re-examined evidence regarding the value of  $\mathbf{v}_\odot$ . We use the value determined by Schönrich, Binney, & Dehnen:

$$\begin{aligned} \mathbf{v}_\odot &= (U_\odot, V_\odot, W_\odot) \\ &= (11.1, 12.24, 7.25) \pm (1, 2, 0.5) \text{ km s}^{-1}, \end{aligned} \quad (8)$$

where  $U_\odot$  is the velocity towards the Galactic Centre,  $V_\odot$  is the velocity in the direction of Galactic rotation,  $W_\odot$  is the velocity perpendicular to the Galactic plane. We quote the systematic uncertainties here, because these dominate the total uncertainty (especially in  $V_\odot$ ).

Many studies that can be used to constrain the value of  $v_0$  are ones which actually constrain the ratio  $v_0/R_0$  – for example using the Oort constants  $A$  and  $B$ , which can be determined locally (e.g. Feast & Whitelock 1997) where  $A - B = v_0/R_0$ . In this study, we use the proper motion of Sgr A\* in the plane of the Galaxy, as determined by Reid & Brunthaler (2004)

$$\mu_{\text{SgrA}^*} = -6.379 \pm 0.026 \text{ mas yr}^{-1}. \quad (9)$$

Since Sgr A\* is expected to be fixed at the Galactic Centre to within  $\sim 1 \text{ km s}^{-1}$ , this proper motion is thought to be almost entirely due to the motion of the Sun around the Galactic Centre  $(v_0 + V_\odot)/R_0$ . This measurement is sufficiently accurate (and the assumed velocity of Sgr A\* sufficiently small) that the greatest uncertainty in the value of  $v_0/R_0$  required for our models actually comes from the uncertainty in the value of  $V_\odot$ !

This measured proper motion is inconsistent with the value of  $A - B$  found by Feast & Whitelock (1997) – who give the most commonly used values for  $A$  and  $B$ . Therefore we do not attempt to use the Oort constants as constraints.

#### 3.2 Terminal velocity curves

For circular orbits in an axisymmetric potential, the peak velocity of the interstellar medium (ISM) along any line-of-sight at Galactic coordinates  $b = 0$  for  $-90 < l < 90$  corresponds to the gas at the tangent point, at Galactocentric radius  $R = R_0 \sin l$ . This is known as the terminal velocity, and is given by

$$v_{\text{term}} = v_c(R_0 \sin l) - v_c(R_0) \sin l. \quad (10)$$

Malhotra (1994, 1995) produced data which gave this peak velocity for a range of lines of sight in the Galaxy from a number of surveys (Weaver & Williams 1974; Kerr et al. 1986; Bania & Lockman 1984; Knapp et al. 1985).

To constrain our model, we have to take account of the effects that non-axisymmetric structure in the Galaxy and non-circular motion of the ISM will have on these data. To do so we follow DB98 in allowing for an uncertainty of  $7 \text{ km s}^{-1}$  in  $v_{\text{term}}$  at any given  $l$ , and restricting ourselves to data at  $|\sin l| > 0.5$ , which is unlikely to be significantly affected by the bar.

#### 3.3 Maser observations

In recent years it has become possible to use Galactic maser sources as targets for astrometric measurements that are sufficiently accurate for parallaxes with uncertainties  $\sim 10 \mu\text{as}$  to be determined for them (e.g. Reid et al. 2009). McMillan & Binney (2010) showed that these observations were consistent with models that placed the masers on near circular orbits with  $v_0/R_0$  similar to the value implied by the proper motion of Sgr A\* (Section 3.1). However  $R_0$  by itself (or, equivalently,  $v_0$ ) was shown to depend quite strongly on the shape of the rotation curve. We constrain our mass model using these data by applying a version of the likelihood analysis used by McMillan & Binney (2010).

The likelihood analysis for these data requires integrating over heliocentric radius the probability density of the observations (parallax, proper motion and radial velocity) given a model for the expected velocity distribution. The maser sources are associated with high mass star formation regions, and are expected to be on near-circular orbits. We model the velocity distribution as circular rotation (with a velocity which depends on the potential), plus a random component that has a Gaussian probability distribution, uniform in direction, with width  $\Delta_v$ . The interested reader can find details in McMillan & Binney (2010). However the analysis in this study differs from that of McMillan & Binney (2010) in that (a) the rotation curve, including  $v_0$ , is defined by the mass model; (b) the value of  $\mathbf{v}_\odot$  is assumed to be that in eq. (8); (c) we include data from Rygl et al. (2010) and Sato et al. (2010), which were not available when the McMillan & Binney (2010) study was performed; (d) in the interests of simplicity we fix  $\Delta_v = 7 \text{ km s}^{-1}$ , close to the best-fitting values found when  $\Delta_v$  was allowed to vary in the previous study, and (e) we assume the masers have no *systematic* velocity offset from the rotation curve – one of the conclusions of McMillan & Binney (2010) was that no such offset is required to explain the data.

### 3.4 Vertical force

Kuijken & Gilmore (1991) used observations of K stars to find the vertical force at 1.1 kpc above the plane at the Solar radius,  $K_{z,1.1}$ . We adopt the value they found

$$K_{z,1.1} = 2\pi G \times (71 \pm 6) M_{\odot} \text{ pc}^{-2} \quad (11)$$

as a constraint.

### 3.5 Observational constraints on the structure at large radii

It is extremely difficult to gain useful constraints on the structure of the Milky Way at large radii from observational data. Any population of dynamical tracers suffers from small number statistics and/or poor observational constraints (especially on proper motion) and because of uncertainty over the associated distribution function, especially the velocity anisotropy, and whether or not all of the population is in fact bound to the Milky Way. This leads to total halo mass estimates which have fractional uncertainties of order unity, e.g.  $1.9^{+3.6}_{-1.7} \times 10^{12} M_{\odot}$  (Wilkinson & Evans 1999),  $0.3$  to  $2.5 \times 10^{12} M_{\odot}$  (Battaglia et al. 2006).

It is possible to take cosmologically motivated simulations and use them to provide some insight into the expected distribution function of a tracer population. Xue et al. (2008) did this for a sample of blue horizontal branch (BHB) stars and found that the mass at radii less than 60 kpc, under this assumption, was  $M(< 60 \text{ kpc}) = (4.0 \pm 0.7) \times 10^{11} M_{\odot}$ . However this value is entirely dependent on the assumption that these cosmological simulations (and their prescriptions for star-formation, feedback etc.) can be relied upon to predict accurately the distribution function of BHB stars.

Another possible source of information is the motion of the Magellanic clouds, but this can only be used straightforwardly if they can be assumed to be bound to the Milky Way – Besla et al. (2007) have shown that this may not be a valid assumption.

Other authors (notably Smith et al. 2007) have attempted to determine the local escape velocity from the Milky Way, and use this to constrain the mass distribution at large radii. However this is determined under the assumption that the population of stars close to the escape velocity can be described by a steady-state distribution function that is depends on energy alone. Stars that have velocities that lie close to the escape velocity are on orbits with very long periods, which means that there is no reason to expect them to be in a steady state, especially since it is likely that a large fraction of halo stars are associated with streams. The assumed distribution function is isotropic (because it only depends on energy), and this assumption is unlikely to be valid – halo stars in cosmological simulations typically have a distribution function which is radially anisotropic.

The ‘‘Timing Argument’’ has been used to constrain the mass of the local group or that of the Milky Way itself. In its simplest form, this assumes that the currently observed radial velocity of M31 towards the Milky Way must be due to the gravitational interaction of the two galaxies overcoming the overall cosmic expansion, and that the corresponding orbital motion must have happened within the age of the universe, and shows that this implies a lower limit on the mass of the two galaxies. Li & White (2008) extended this

argument by searching cosmological simulations for comparable galaxy pairs that were moving towards one another in the simulations. They also looked for galaxy pairs comparable to the Milky Way and Leo I, which are moving away from one another, and used the Timing Argument under the assumption that Leo I is moving towards apocentre for a second time. This yielded a probability distribution on  $M_v$  with lower and upper quartiles  $[1.78, 3.09] \times 10^{12} M_{\odot}$  and a 95% confidence lower limit of  $7.94 \times 10^{11} M_{\odot}$ . This approach rather relies upon Leo I being bound to the Milky Way, which it may not be, and the authors note that, compared to the Milky Way/M31 system, the ‘‘complex dynamical situation offers greater scope for uncertainty’’.

While Wilkinson & Evans (1999) quoted a very uncertain figure for the total halo mass, they also found that the mass within 50 kpc was a more robustly determined quantity:  $M_{50} = 5.4^{+0.2}_{-3.2} \times 10^{11} M_{\odot}$ . The quoted uncertainty is very asymmetrically distributed about the most likely value, and we choose to see this result as providing an upper bound on the mass inside 50 kpc, which we take into account in our analysis through the probability distribution

$$P(M_{50}) = \begin{cases} C & \text{for } M_{50} \leq M_{\text{WE}} \\ C \exp\left(-\left[\frac{M_{50}-M_{\text{WE}}}{\delta_{M_{50}}}\right]^2\right) & \text{for } M_{50} > M_{\text{WE}} \end{cases} \quad (12)$$

where  $M_{\text{WE}} = 5.4 \times 10^{11} M_{\odot}$ ,  $\delta_{M_{50}} = 2 \times 10^{10} M_{\odot}$ , and  $C$  is a normalisation constant. The only effect of using this  $P(M_{50})$  constraint is to penalise models with  $M_{50} > M_{\text{WE}}$ . For models with  $M_{50} \ll M_{\text{WE}}$  this would be an inappropriate choice but, as we show in Section 5, that is not the case for this study.

Given the difficulties described here, we have decided not to constrain our model with the results of any other study. We do compare our best-fitting model to them (Section 6)

## 4 FITTING THE MODELS

We use Bayesian statistics to find the probability density function (pdf) for our model parameters given the kinematic data described in Section 3, and the prior probabilities given in Section 2 (and the value of  $R_0$ , given in Section 3). We refer to the parameters collectively as  $\theta$ , and the data as  $d$ . Bayes theorem tells us that this pdf is then

$$p(\theta|d) = \frac{\mathcal{L}(d|\theta) p(\theta)}{p(d)} \quad (13)$$

where the total likelihood  $\mathcal{L}(d|\theta)$  is the product of the likelihoods associated with each kinematic data-set or constraint described in Section 3, and  $p(\theta)$  is the probability of a parameter set given the prior probability distributions described in Section 2 (and that on  $R_0$ ). The Bayesian Evidence,  $p(d)$ , is often a very important quantity, but in this study it is an unimportant normalisation constant, so we ignore it.

Given the constraints we have applied to the components described in Section 2, we are left with 8 model parameters that we allow to vary: The scale-lengths and density normalisations of the thin and thick discs ( $R_{\text{d,thin}}, \Sigma_{\text{d,0,thin}}, R_{\text{d,thick}}, \Sigma_{\text{d,0,thick}}$ ); the density normalisation – and thus mass – of the bulge ( $\rho_{\text{b,0}}$ ); the scale-length

Property constrained	Constraint	Section described	Source
Bulge profile	see equation (1)	2.1	Bissantz & Gerhard (2002)
$M_b$	$(8.9 \pm 0.89) \times 10^9 M_\odot$	2.1	Bissantz & Gerhard (2002)
Disc profile	Double exponential	2.2	-
$z_{d,thin}$	0.3 kpc	2.2	Jurić et al. (2008)
$z_{d,thick}$	0.9 kpc	2.2	Jurić et al. (2008)
$R_{d,thin}$	$2.6 \pm 0.52$ kpc	2.2	Jurić et al. (2008)
$R_{d,thick}$	$3.6 \pm 0.72$ kpc	2.2	Jurić et al. (2008)
$f_{d,\odot}$	$0.12 \pm 0.012$	2.2	Jurić et al. (2008)
Halo profile	NFW profile	2.3	Navarro et al. (1996)
$M_*/M_v$	see equation (5)	2.3	Li & White (2009)
$\ln c_v$	$2.256 \pm 0.272$	2.3	Boylan-Kolchin et al. (2010)
$R_0$	$8.33 \pm 0.35$ kpc	3.1	Gillessen et al. (2009)
$\mu_{SgrA^*}$	$-6.379 \pm 0.026$ mas yr $^{-1}$	3.1	Reid & Brunthaler (2004)
$K_{z,1.1}$	$2\pi G \times (71 \pm 6) M_\odot \text{ pc}^{-2}$	3.4	Kuijken & Gilmore (1991)
$M_{50}$	$\lesssim 5.4 \times 10^{11} M_\odot$ , see eq. (12)	3.5	Wilkinson & Evans (1999)
Kinematic data		Section described	Source
Terminal velocities		3.2	Malhotra (1994, 1995)
Masers observations		3.3	Reid et al. (2009); Rygl et al. (2010); Sato et al. (2010)

**Table 1.** Summary of all the constraints applied to the models.

and density normalisation of the CDM halo ( $r_h, \rho_{h,0}$ ); and the solar radius ( $R_0$ ). While each of these parameters is free to vary, each one is explicitly or implicitly associated with at least one prior probability distribution.

To explore the pdf  $p(\theta|d)$  we use the Metropolis algorithm (Metropolis et al. 1953), which is a Markov Chain Monte Carlo method for drawing a representative sample from a probability distribution. This allows us both to find the peak of the pdf (to reasonable accuracy) and to characterise its shape. We start with some choice for the parameters  $\theta_n$ , and calculate  $p(\theta_n|d)$ . We then

- (i) choose a trial parameter set  $\theta'$  by moving from  $\theta_n$  in all directions in parameter space, by an amount chosen at random from a “proposal density”  $Q(\theta', \theta_n)$  (see below);
- (ii) determine  $p(\theta'|d)$ ;
- (iii) choose a random variable  $r$  from a uniform distribution in the range  $[0,1]$ ;
- (iv) if  $p(\theta'|d)/p(\theta_n|d) > r$ , accept the trial parameter set, and set  $\theta_{n+1} = \theta'$ . Otherwise set  $\theta_{n+1} = \theta_n$ .
- (v) Return to step (i), replacing  $\theta_n$  with  $\theta_{n+1}$ .

The first few values of  $\theta$  are ignored as “burn-in”, which helps to remove the dependence on the initial value of  $\theta$ .

The pdf is quite narrow in some directions in the parameter space, but these directions are not simply parallel to the coordinate axes. Therefore it is efficient to use a proposal density which is aligned (to a reasonable approximation) with the pdf – this is allowed by the Metropolis algorithm as the only constraints on  $Q(\theta_1, \theta_2)$  are that it is symmetrical with respect to swapping  $\theta_1$  and  $\theta_2$ , and makes it possible to reach any point in phase space. We therefore use the Metropolis algorithm in two phases – first with a proposal density which is simply made up of Gaussian distributions in each parameter individually, with associated step size chosen by hand. The resultant chain (of a relatively short length, and after a burn-in) is then used to construct a covariance matrix that is then used to define a new  $Q(\theta', \theta_n)$ . This new proposal density is, again, a multi-variate Gaussian but with the principal axes now aligned

to the eigenvectors of the covariance matrix, and with the step size in each direction related to the respective eigenvalues. The chain produced with this second proposal density is then used in all calculations.

## 5 RESULTS

In Figures 1 to 3 we plot the probability density functions of various quantities associated with our models, marginalised over all parameters, as determined by the Metropolis algorithm. Where appropriate we also plot the prior probability distribution directly associated with each. The value associated with our best-fitting model is indicated on each plot with a dashed vertical line. In Table 2 we give the parameters of our best-fitting model, and those of what we call the “convenient” model (see below), along with the mean and standard deviation for each parameter, marginalised over other parameters in the pdf.

The marginalised distributions do not give a sense of the correlations between parameters, so in Table 3 we show the correlation matrix of the various parameters  $\theta$ . For the  $i,j$ th component,  $\text{corr}(\theta_i, \theta_j)$ , this takes the value

$$\text{corr}(\theta_i, \theta_j) = \frac{\text{cov}(\theta_i, \theta_j)}{\sigma_i \sigma_j} \quad (14)$$

where  $\text{cov}(\theta_i, \theta_j)$  is the covariance. A value of 1 corresponds to a perfect correlation,  $-1$  corresponds to a perfect anti-correlation, and 0 corresponds to no correlation. The correlation matrix is manifestly symmetric, so in Table 3 we only show half of it.

The strongest correlations or anti-correlations are between parameter pairs that describe a single component – this explains, for example, why the spread in  $M_*$  is so much smaller than the spread in  $\Sigma_{d,0,thin}$  or  $\Sigma_{d,0,thick}$ , and why the standard deviation of  $\rho_{h,0}$  can be as large as  $\sim 50\%$  while that in  $M_v$  is  $\sim 20\%$  (and that in  $M_{50}$  is even smaller, see below). Other fairly strong correlations are noticeable

	$\Sigma_{d,0,\text{thin}}$	$R_{d,\text{thin}}$	$\Sigma_{d,0,\text{thick}}$	$R_{d,\text{thick}}$	$\rho_{b,0}$	$\rho_{h,0}$	$r_h$	$R_0$		
Best	816.6	2.90	209.5	3.31	95.6	0.00846	20.2	8.29		
Convenient	753.0	3	182.0	3.5	94.1	0.0125	17	8.5		
Mean	741	3.00	238	3.29	95.5	0.012	18.0	8.29		
Std. Dev.	123	0.22	110	0.56	6.9	0.006	4.3	0.16		
	$v_0$	$M_b$	$M_*$	$M_v$	$K_{z,1.1}$	$\Sigma_{d,\odot}$	$\rho_{*,\odot}$	$\rho_{h,\odot}$	$f_{d,\odot}$	$g_{*,\odot}/g_{h,\odot}$
Best	239.1	8.97	66.1	1400	77.7	63.9	0.087	0.0104	0.122	1.38
Convenient	244.5	8.84	65.1	1340	75.4	60.3	0.083	0.0111	0.121	1.12
Mean	239.2	8.96	64.3	1260	76.5	62.0	0.085	0.0106	0.120	1.29
Std. Dev.	4.8	0.65	6.3	240	5.3	7.6	0.010	0.0010	0.012	0.30

**Table 2.** Parameters (upper) and derived properties (lower) of our best-fitting model, our “convenient” model, and mean and standard deviation marginalised over all the (other) parameters in the pdf. The density profile of the bulge follows the description in eq. (1), with other parameters as described directly below that equation; the density profiles of the discs follow eq. (3), with scale-heights 0.3 and 0.9 kpc for thin and thick discs respectively; and the halo density profile is that given in eq. (4).  $\Sigma_{d,\odot}$  is the disc surface density at the Sun;  $\rho_{*,\odot}$  and  $\rho_{h,\odot}$  are the local densities of the stellar component and CDM halo, respectively;  $g_{*,\odot}/g_{h,\odot}$  is the ratio of the gravitational force on the Sun from the stellar component ( $g_{*,\odot}$ ) to that from the CDM halo ( $g_{h,\odot}$ ) – this is a measure of whether the Galaxy is disc- or halo-dominated. The means and standard deviations of the parameters are not always particularly helpful statistics, as they say nothing about the correlations between parameters. It should also be noted that the true uncertainties on  $\Sigma_{d,0,\text{thin}}$ ,  $\Sigma_{d,0,\text{thick}}$  and  $\rho_{*,\odot}$  are considerably larger than the standard deviations quoted here, because the disc scale heights have been held constant (see Section 5.1). Distances are quoted in units of kpc, velocity in  $\text{km s}^{-1}$ , masses in  $10^9 M_\odot$ , surface densities in  $M_\odot \text{pc}^{-2}$ , densities in  $M_\odot \text{pc}^{-3}$ , and  $K_{z,1.1}$  in units of  $(2\pi G) \times M_\odot \text{pc}^{-2}$ .

	$\Sigma_{d,0,\text{thin}}$	$R_{d,\text{thin}}$	$\Sigma_{d,0,\text{thick}}$	$R_{d,\text{thick}}$	$\rho_{b,0}$	$\rho_{h,0}$	$r_h$	$R_0$
$\Sigma_{d,0,\text{thin}}$	1							
$R_{d,\text{thin}}$	−0.727	1						
$\Sigma_{d,0,\text{thick}}$	−0.467	0.483	1					
$R_{d,\text{thick}}$	0.468	−0.304	−0.845	1				
$\rho_{b,0}$	−0.197	0.167	−0.047	0.037	1			
$\rho_{h,0}$	−0.632	0.328	−0.247	0.151	−0.009	1		
$r_h$	0.588	−0.299	0.221	−0.130	0.003	−0.877	1	
$R_0$	−0.369	0.602	−0.071	0.157	−0.017	0.373	−0.338	1

**Table 3.** Correlation matrix for the model parameters. A value of 1 corresponds to perfect correlation (e.g. any parameter with itself), −1 corresponds to perfect anti-correlation, 0 to no correlation. The strongest relationships are anti-correlations between parameters which, taken in combination, define the mass of a given component or its density at a given point – i.e. between  $\rho_{h,0}$  &  $r_h$ , between  $\Sigma_{d,0,\text{thin}}$  &  $R_{d,\text{thin}}$  and between  $\Sigma_{d,0,\text{thick}}$  &  $R_{d,\text{thick}}$ .

between  $\Sigma_{d,0,\text{thin}}$  and virtually every other parameter, and between  $R_{d,\text{thin}}$  and  $R_0$ .

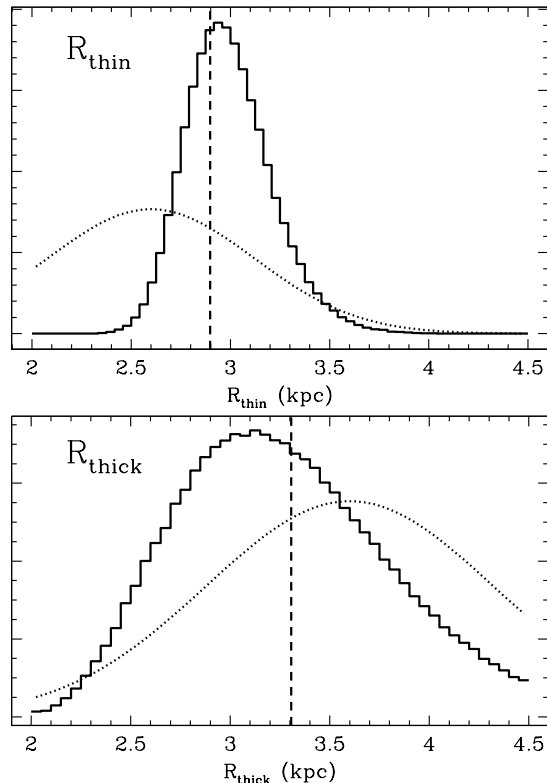
Figure 1 shows the pdfs associated with the disc scale-lengths,  $R_{\text{thin}}$  and  $R_{\text{thick}}$ . Both have peaks around 3 kpc, and the best-fitting values also lie quite near 3 kpc. That the kinematic data (primarily the terminal velocity curve) drives the two scale-lengths towards a similar value is unsurprising – the two discs have effects on the dynamics of the Galaxy that scale very similarly. The posterior pdf on  $R_{\text{thin}}$  is much narrower than the prior, which is not true to the same extent for  $R_{\text{thick}}$ ; this is not surprising given that the thin disc is more massive than the thick disc so the constraints based on kinematic data are more likely to dominate the photometric constraints in the thin disc pdf.

DB98 considered mass models with four different scale-lengths between 2 kpc and 3.2 kpc, and halo density profiles which could take on a wide range of power-law slopes in their inner and outer parts. They found that the models with disc scale-lengths of 2.4 kpc or smaller required haloes which, unrealistically, had density profiles which *increased* with radius ( $\rho \propto r^2$ ) in their inner parts, whereas the models with scale-lengths 2.8 kpc or greater did not. Therefore, given that we require a halo profile that increases as  $\rho \propto r^{-1}$

in its centre, we should not be surprised that this favours models with disc scale-lengths larger than 2.5 kpc.

Figure 2 shows the posterior pdfs for the position of the Sun and the circular speed at the Sun in our models. The posterior pdf on  $R_0$  is much narrower than the prior, with a standard deviation of 0.16 kpc, compared to an uncertainty on the prior of 0.33 kpc. The posterior pdf on  $v_0/R_0$  is of a similar width to that which would be seen if our only information on its value was the proper motion of Sgr A\*, but it is slightly displaced to higher values of  $v_0/R_0$ . This displacement is most likely due to the maser data (Section 3.3), which McMillan & Binney (2010) showed are best fit by models in which  $v_0/R_0$  is slightly larger than the value implied by equation (9). The corresponding value of  $v_0$ ,  $239.2 \pm 4.8 \text{ km s}^{-1}$ , has a significantly smaller uncertainty than simply combining the Gillessen et al. (2009) value of  $R_0$  and the Reid & Brunthaler (2004) value of  $\mu_{\text{SgrA}^*}$  gives, primarily because of the tighter constraints on  $R_0$ .

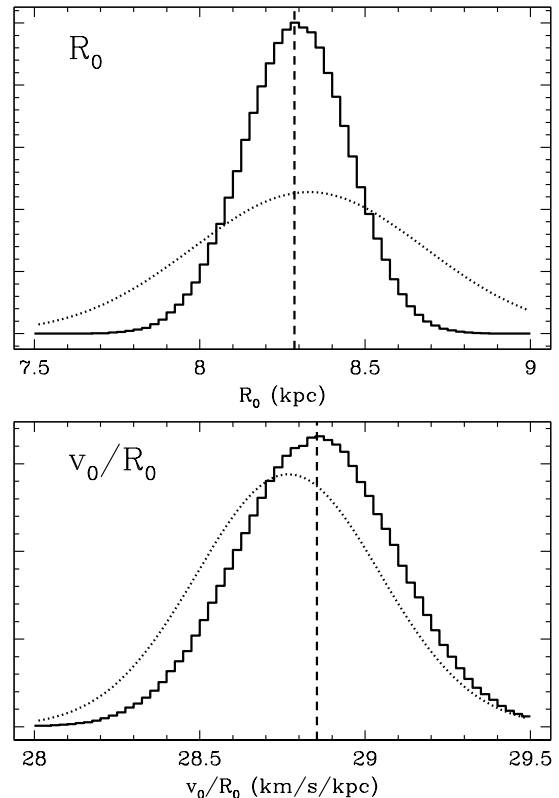
The posterior pdf on bulge masses (Figure 3, top left) is close to our prior from Bissantz & Gerhard (2002). The bottom right panel of Figure 3 shows that the stellar mass of the models is tightly peaked at values towards the high end of those we would expect given the virial mass, and



**Figure 1.** Histogram of the pdf of the thin (top) and thick (bottom) disc scale-lengths in our models (solid histogram) normalised over all other parameters, compared in each case to the prior pdf described in Section 2.2 (dotted). In each plot the value corresponding to our best-fitting model is shown as a dashed vertical line. In each case the posterior pdf is peaked near to 3 kpc.

our constraint on this relationship given in equation (5). This can equivalently be thought of as corresponding to low halo masses, given the stellar masses. Compared to the prior pdfs taken from theoretical considerations (Section 2.3), the haloes have low masses and high concentrations ( $\ln c_{v'} = 2.73$  for the best-fitting model, and with mean and standard deviation 2.83 and 0.18 respectively). The halo is likely to be sub-dominant at  $R_0$  in the sense that the gravitational force on the Sun is mostly from the stellar component, though this is not the case for all models (ratio of gravitational forces  $g_{*,\odot}/g_{h,\odot} = 1.29 \pm 0.30$ ). The mass inside 50 kpc of our best-fitting model is  $5.3 \times 10^{11} M_\odot$ , with mean and standard deviation  $5.1 \pm 0.4 \times 10^{11} M_\odot$ . This is close to the mass limit we take from Wilkinson & Evans (1999), indicating that the latter is an important upper limit on our models.

The local density normalisation  $f_{d,\odot}$  is 0.122 for our best-fitting model. The mean and standard deviation are  $0.120 \pm 0.012$ , which is exactly what one would expect from our prior alone. This is due to the fact that our kinematic data does not provide us with any great ability to discriminate between mass in either of the two discs. Therefore, the value of  $f_{d,\odot}$  almost entirely defines (with the disc scale-heights) our constraints on the relative contributions of the thick and thin disc.

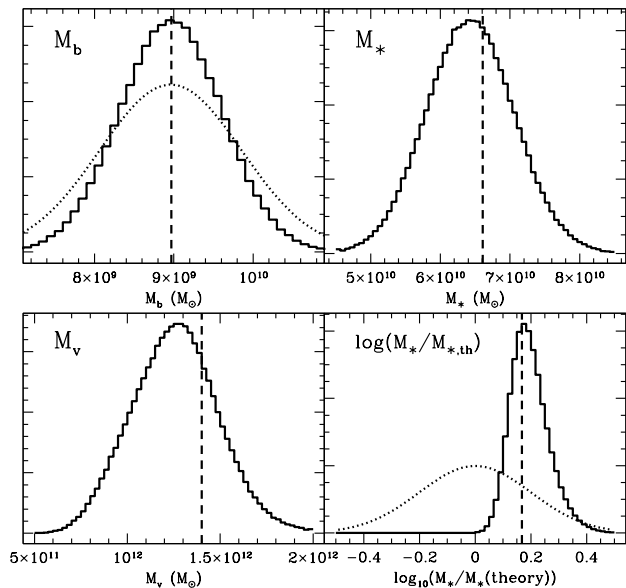


**Figure 2.** Histogram of the pdf of  $R_0$  (top) and  $v_0/R_0$  (bottom) normalised over all parameters (solid line). The prior pdf on  $R_0$  and the pdf on  $v_0/R_0$  associated with the apparent motion of Sgr A\* (Section 3.1) are plotted as dotted lines. As in Figure 1 the value corresponding to our best-fitting model is shown as a dashed vertical line.

Catena & Ullio (2010) used a rather similar method to the one described in this paper with the single intention of determining the local dark matter density, finding a value  $\sim 0.39 \text{ GeV cm}^{-3}$ . We have made a number of different assumptions, and taken into account different constraints, but many of the key assumptions (axisymmetry, exponential disc and a halo profile motivated by CDM simulations) are the same, and we find a similar local dark matter density  $0.40 \pm 0.04 \text{ GeV cm}^{-3}$ . It is well worth noting that the uncertainty we quote here is purely the statistical uncertainty associated with this set of parametrized models, and the constraints we have chosen to apply. In particular, we have made the approximations that the discs are well described by exponential profiles at all radii, and that the halo density profile is well described by a spherically symmetric NFW profile. These assumptions and others, including ones which are not directly related to the dark matter profile, will have a significant effect on the value found for the local dark matter density. We expect that the systematic uncertainties on this value are much larger than the statistical uncertainty we state here.

It is clear from Figures 1 to 3 that there is a wide range of models that fit our constraints almost as well as our best-fitting model. It is often convenient to use models in which certain key parameters are chosen to take simple values (for





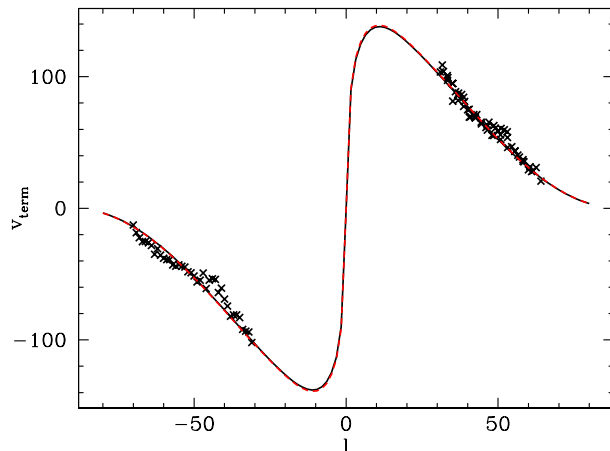
**Figure 3.** Histogram of the pdf for the bulge mass  $M_b$  (top-left), the total stellar mass  $M_*$  (top-right), the virial mass  $M_v$  (bottom-left) and the log of the ratio of  $M_*$  to the value of  $M_*$  predicted by equation (5) (bottom-right). In first and last case, where a straightforward prior applies to the quantity shown, it is plotted as a dotted curve. Again, the value corresponding to our best-fitting model is shown as a dashed vertical line in each plot.

example it is common to take  $R_0 = 8$  or  $8.5$  kpc), and we provide such a model in Table 2. To take account of the correlations between parameters we chose to build this “convenient” model one step at a time. We first hold  $R_0$  constant at a suitable value determined from Figure 2, and find the pdf associated with these models, which we use to choose a suitable thin disc scale length  $R_{d,thin}$ . We then repeat this process to find a suitable value for  $R_{d,thick}$ , and then again for  $r_h$ . Following this procedure we take  $R_0 = 8.5$  kpc,  $R_{d,thin} = 3$  kpc,  $R_{d,thick} = 3.5$  kpc and  $r_h = 17$  kpc.

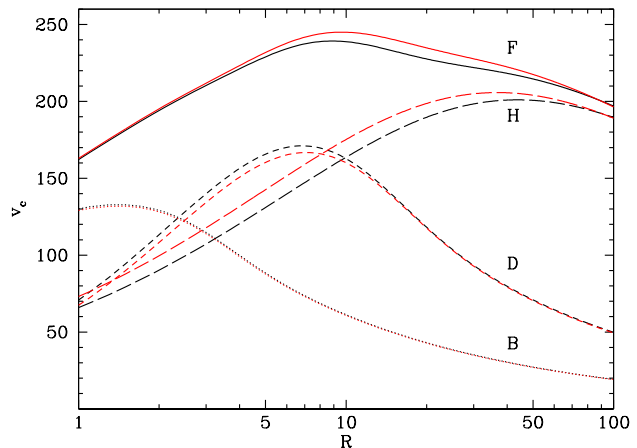
In Figure 4 we compare the terminal velocity curve associated with our best-fitting model, and with the “convenient” model, to the observational data (Section 3.2). The difference between the curves of the two models is very small, and both provide a good fit to these data. In Figure 5 we plot the rotation curves of the best-fitting and convenient models, and the decomposition into the components due to the bulge, discs, and halo. The two models are more clearly distinguished in this plot, primarily because of the larger value of  $R_0$  chosen for the convenient model, which results in a higher value for  $v_0$ , because of the strong correlation between the two.

### 5.1 Effect of changing the disc scale-heights

In all of the models discussed thus far (and shown in Figures 1–5 and Tables 2 & 3) the scale-heights of the discs have been held constant at 300 pc and 900 pc for the thin and thick discs respectively. These values were chosen because they are the best-fitting scale-heights given by Jurić et al.



**Figure 4.** Comparison of the terminal velocity curve predicted by our best-fitting model (solid curve), the “convenient” model (dashed red curve) and the terminal velocity measurements from the Galaxy (crosses). The two model curves are nearly indistinguishable.



**Figure 5.** The rotation curve for our best-fitting (black) and convenient (red) models – note that radius is plotted logarithmically. The solid line, labelled F, is the full rotation curve, with the other curves showing, in each case, the contribution of the bulge (B, dotted), discs (D, short-dashed) and halo (H, long-dashed).

(2008), and were held constant because we do not expect the kinematic data described in Section 3 to strongly constrain the vertical density profile. It is, however, important that we check this assumption by considering different scale-heights for the disc. Therefore we now consider models which have a thin disc scale-height  $h_{d,thin}$  of 250, 300 or 350 pc and a thick disc scale-height  $h_{d,thick}$  of 750, 900 or 1050 pc, in all nine possible combinations.

One change to our prior must be considered because of the effect of changing the scale-height of the disc. The local density normalisation,  $f_{d,\odot} = \rho_{thick}(R_\odot, z_\odot)/\rho_{thin}(R_\odot, z_\odot)$ , is defined at the Sun’s position, so if the scale-heights change, the value of  $f_{d,\odot}$  changes without the surface densities of the two discs changing. It has been noted (by, for example, Reylé & Robin 2001) that there is an anti-correlation between the values of  $f_{d,\odot}$  and  $h_{d,thick}$  found

by studies using star counts, which can indeed be seen in figures 21 & 24 of Jurić et al. (2008). We therefore approximate that  $f_{d,\odot}$  simply changes with  $h_{d,\text{thick}}$ , taking the values  $f_{d,\odot} = 0.14, 0.12, 0.10$  for  $h_{d,\text{thick}} = 750, 900, 1050$  pc respectively, with an uncertainty in  $f_{d,\odot}$  of  $\pm 0.012$  in all cases (as previously). These values are taken by eye from figure 21 of Jurić et al. (2008), with a correction for the recognised biases in the data. We ignore any possible correlation between  $f_{d,\odot}$  and  $h_{d,\text{thin}}$  – an anti-correlation of this type is seen in figures 21 & 24 of Jurić et al. (2008), but could easily be related to the relationship between  $f_{d,\odot}$  and  $h_{d,\text{thick}}$ , as the values they derive for  $h_{d,\text{thin}}$  and  $h_{d,\text{thick}}$  are also correlated.

We give the mean and standard deviation for the parameters (and derived quantities) of all these models in the appendix. Here we discuss the most significant findings.

Most importantly, the changes in scale heights have very little impact on the overall structure of the Galaxy models. The bulge mass and virial mass of the Galaxy are virtually unchanged, and the total stellar mass changes by at most  $\sim 5\%$ , which is much less than the standard deviation of the value. The disc scale lengths are also barely changed (changes of at most  $\sim 2\%$  in  $R_{d,\text{thin}}$  and  $\sim 4\%$  in  $R_{d,\text{thick}}$ ). The surface density at the Sun is also largely unchanged, with changes of  $\sim 1\%$  in  $K_{z,1.1}$  and the local disc surface density  $\Sigma_{d,\odot}$  changing by  $\sim 5\%$ .

The only significant change is in the values of  $\Sigma_{d,0,\text{thin}}$  and  $\Sigma_{d,0,\text{thick}}$ , which show a transfer of mass from the thick disc to the thin disc as  $h_{d,\text{thin}}$  increases. The mean value of  $\Sigma_{d,0,\text{thin}}$  rises from  $\sim 700 \text{ M}_\odot \text{ pc}^{-2}$  for models with  $h_{d,\text{thin}} = 0.25$  to  $\sim 800 \text{ M}_\odot \text{ pc}^{-2}$  for those with  $h_{d,\text{thin}} = 0.35$ , while the value of  $\Sigma_{d,0,\text{thick}}$  declines from  $\sim 280 \text{ M}_\odot \text{ pc}^{-2}$  to  $\sim 210 \text{ M}_\odot \text{ pc}^{-2}$  over the same range. This then affects the fraction of the total disc mass found in each disc, and the value of  $f_{d,\odot}$ .

In every case the posterior pdf of  $f_{d,\odot}$  is almost precisely the same as the prior. This indicates that the prior on  $f_{d,\odot}$  is, in all cases, the dominant factor determining the how the Galaxy model’s total disc mass is divided between the two discs.

We can therefore see that holding the disc scale-heights at fixed values to produce the statistics in Table 2 does not significantly alter the quoted standard deviations, except for parameters related to the division of the disc material into the thick and thin discs. Here the true uncertainties are significantly larger than the quoted values because the uncertainty in scale-heights must be taken into account.

## 6 COMPARISON TO OTHER STUDIES

The thin disc scale length of our best-fitting model is at the larger end of the Jurić et al. (2008) range, which we used as a prior. It is therefore also larger than the values in the range 2 to 2.5 kpc found by some other recent studies (Ojha et al. 1999; Chen et al. 1999; Siegel et al. 2002), and the value used for the old (age  $> 0.15$  Gyr) disc in the Besançon Galaxy model (2.53 kpc, Robin et al. 2003). On the other hand it lies close to the scale length found by Kent, Dame, & Fazio (1991),  $3 \pm 0.5$  kpc, and towards the lower end of the range found by López-Corredoira et al. (2002)  $3.3^{+0.5}_{-0.4}$  kpc. The stellar mass of our best-fitting model is  $6.61 \times 10^{10} \text{ M}_\odot$  (with mean and standard de-

viation  $6.43 \pm 0.63 \times 10^{10}$ ). This compares to the value  $6.1 \pm 0.5 \times 10^{10} \text{ M}_\odot$  found as a “back of the envelope” estimate by Flynn et al. (2006). The local disc surface density ( $\Sigma_{d,\odot} = 63.9 \text{ M}_\odot \text{ pc}^{-2}$  best-fitting;  $62.0 \pm 7.6 \text{ M}_\odot \text{ pc}^{-2}$  mean  $\pm$  standard deviation) is somewhat larger than suggested by studies which count identified matter in the Solar neighbourhood, such as that of Kuijken & Gilmore (1989) who found  $\Sigma_{d,\odot} = 48 \pm 8 \text{ M}_\odot \text{ pc}^{-2}$ , or Flynn et al. (2006) who found  $\Sigma_{d,\odot} \sim 49 \text{ M}_\odot \text{ pc}^{-2}$ .

Our best-fitting bulge mass is close to the centre of the prior distribution we take from Bissantz & Gerhard (2002). This is significantly larger than most of the values found by DB98, and comparable to the bulge masses found by Widrow et al. (2008). This is somewhat smaller than the masses determined by many purely photometric studies (e.g. Dwek et al. 1995; Picaud & Robin 2004; Launhardt et al. 2002), but, as noted in Section 2.1, there is an important dependence on assumptions made about the disc. We note that the stellar mass within the inner 3 kpc in our best-fitting model,  $\sim 2.4 \times 10^{10} \text{ M}_\odot$ , is very close to the bulge mass found by Picaud & Robin (2004) assuming a central hole in the disc:  $2.4 \pm 0.6 \times 10^{10} \text{ M}_\odot$ .

The baryonic Tully-Fisher relation describes the relationship between total baryonic mass  $M_{\text{baryon}}$  and the circular speed  $V_c$  at some radius, observed in external galaxies. McGaugh et al. (2010) found  $\log_{10} M_{\text{baryon}} = 4.0 \log_{10} V_c + 1.65$  for disc galaxies, with a scatter that is entirely consistent with observational uncertainty, where  $1.1 \times V_c$  was the actual observed circular speed. Our models have  $\log_{10} M_{\text{baryon}} - (4.0 \log_{10} V_c + 1.65) = -0.19 \pm 0.05$ , suggesting that the Milky Way has a significantly higher rotational speed (or, equivalently, lower baryonic mass) than the Tully-Fisher relation predicts. A similar offset from the baryonic Tully-Fisher relationship was found for the Milky Way by Flynn et al. (2006).

In Section 3.5, we noted the difficulty of finding rigorous constraints on Galactic structure at large radii, and described some of problems with commonly cited constraints. It is still instructive to compare our models to the results of these previous studies. The escape speed at the Sun of our best-fitting model is  $622 \text{ km s}^{-1}$ , with mean and standard deviation over all models of  $606 \text{ km s}^{-1}$  and  $26 \text{ km s}^{-1}$  respectively. This compares to the quoted 90% confidence interval 498 to  $608 \text{ km s}^{-1}$  of Smith et al. (2007). The mass inside 60 kpc is  $6.2 \times 10^{11} \text{ M}_\odot$  (best-fitting), with mean and standard deviation  $5.9 \pm 0.5 \times 10^{11} \text{ M}_\odot$ . This is rather larger than the Xue et al. (2008) value of  $4.0 \pm 0.7 \times 10^{11} \text{ M}_\odot$ . Xue et al. used cosmological simulations to predict the velocity anisotropy of the BHB population they were studying – this leads to predictions of rather high radial anisotropy, which drives the mass estimate towards lower values than more isotropic or tangentially anisotropic velocity distributions would suggest. DB98 adopted the constraint on  $M_{100}$ , the mass inside 100 kpc,  $M_{100} = 7 \pm 2.5 \times 10^{11} \text{ M}_\odot$ , based on then-available data. Our best-fitting model ( $M_{100} = 9.0 \times 10^{11} \text{ M}_\odot$ ) and the ensemble of models as a whole (mean  $\pm$  standard deviation:  $8.4 \pm 0.9 \times 10^{11}$ ) fit well within this range. The virial mass of our best-fitting model  $M_v = 1.40 \times 10^{12} \text{ M}_\odot$  (mean  $\pm$  standard deviation:  $1.26 \pm 0.24 \times 10^{12} \text{ M}_\odot$ ) lies well within the ranges quoted by Wilkinson & Evans (1999) or Battaglia et al. (2006), and slightly below the

lower quartile of the pdf given by Li & White (2008), but well above their 95% confidence limit.

## 7 CONCLUSIONS

We have presented a simple Bayesian method for applying photometrically and kinematically derived constraints, and theoretical understanding of galaxy structure, to parametrized axisymmetric models of the Galaxy in order to investigate the distribution of mass in its various components. We have applied this method to models with an axisymmetric bulge, exponential discs and an NFW halo.

The method we have described is sufficiently general that it could be applied to any sensible parametrized axisymmetric mass model, with a wide range of kinematic, photometric, theoretical or other constraints. The specific constraints we apply are summarised in Table 1. We have shown that these constraints still allow a wide range of Galaxy models, and have found a best-fitting model, as well as a model that is best-fitting after key parameters have been fixed at convenient values. These models will provide a suitable starting point for producing fully dynamical Galaxy models.

We have also shown that the main features of our models are unchanged when we consider different disc scale-heights, except to the extent that they alter our prior probability distribution on the relative contributions of the thin and thick discs, which alters their relative contributions in the models. It should therefore be noted that the (already weak) constraints we have on the *relative* contributions of the two discs are even weaker when the uncertainty in disc scale-heights is taken into account. This constraint on the ratio of thick to thin disc contributions is almost entirely that from our Jurić et al. (2008) prior, and a different choice of prior could result in a significantly different ratio.

The kinematic data we consider here does not help us to constrain the vertical density profile of the Galactic discs, but it is clear that kinematic data can be used in combination with star counts to provide greater insight into the Galactic potential above the plane (e.g. Burnett 2010), which could then be used to improve these models.

Applying the need for consistency in our model allows us to find tighter constraints on individual parameters than those we begin with. This is particularly noticeable in the constraints our posterior pdf place on the thin disc scale-length ( $R_{d,thin}$ ), the Solar radius ( $R_0$ ), and the circular velocity at the Solar radius ( $v_0$ ). We find that the Galaxy's dark-matter halo concentration,  $c_v$ , is larger than the average value predicted by simulations. The Galaxy's halo is less massive than the expected value from cosmological simulations, given the Galaxy's stellar mass (or, equivalently, the stellar mass is higher than would be expected). In contrast, the stellar mass of the Milky Way is lower than the baryonic Tully-Fisher relation would suggest given its circular velocity – the discrepancy between the two expectations for the stellar mass is related to the high concentration of the halo. In addition to the uncertainty on individual parameters, we are able to find the correlations between different parameters demanded by our constraints.

The results described in this paper are all dependent on the choice of parametrized model and applied constraints. In

particular we have not attempted to take account of the effect on the CDM halo of baryonic processes, or to consider a CDM halo which is not spherically symmetric. The systematic uncertainties on the quantities we describe are almost undoubtedly larger than the quoted statistical uncertainties.

## ACKNOWLEDGMENTS

I am very grateful to James Binney for insightful comments, and a careful reading of a draft of this paper. I thank other members of the Oxford dynamics group, both past and present, for valuable discussions. This work is supported by a grant from the Science and Technology Facilities Council.

## REFERENCES

- Abazajian K. N., Adelman-McCarthy J. K., Agüeros M. A., et al., 2009, *ApJS*, 182, 543
- Allgood B., Flores R. A., Primack J. R., Kravtsov A. V., Wechsler R. H., Faltenbacher A., Bullock J. S., 2006, *MNRAS*, 367, 1781
- Bania T. M., Lockman F. J., 1984, *ApJS*, 54, 513
- Battaglia G., Helmi A., Morrison H., Harding P., Olszewski E. W., Mateo M., Freeman K. C., Norris J., Shectman S. A., 2006, *MNRAS*, 370, 1055
- Bell E. F., Zucker D. B., Belokurov V., et al., 2008, *ApJ*, 680, 295
- Belokurov V., Zucker D. B., Evans N. W., et al., 2006, *ApJL*, 642, L137
- Besla G., Kallivayalil N., Hernquist L., Robertson B., Cox T. J., van der Marel R. P., Alcock C., 2007, *ApJ*, 668, 949
- Binney J., 2010, *MNRAS*, 401, 2318
- Binney J., Gerhard O., Spergel D., 1997, *MNRAS*, 288, 365
- Binney J. J., McMillan P. J., 2011, *MNRAS*, accepted (arXiv:1101.0747)
- Bissantz N., Gerhard O., 2002, *MNRAS*, 330, 591
- Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., 2010, *MNRAS*, 406, 896
- Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, *MNRAS*, 398, 1150
- Burnett B. C. M., 2010, Ph.D. Thesis, University of Oxford
- Caldwell J. A. R., Ostriker J. P., 1981, *ApJ*, 251, 61
- Catena R., Ullio P., 2010, *JCAP*, 8, 4
- Chen B., Figueras F., Torra J., Jordi C., Luri X., Galadí-Enríquez D., 1999, *A&A*, 352, 459
- de Blok W. J. G., 2010, *Advances in Astronomy*, 2010
- Debattista V. P., Moore B., Quinn T., Kazantzidis S., Maas R., Mayer L., Read J., Stadel J., 2008, *ApJ*, 681, 1076
- Dehnen W., Binney J., 1998, *MNRAS*, 294, 429
- Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., Battye R. A., Booth C. M., 2010, *MNRAS*, 405, 2161
- Dwek E., Arendt R. G., Hauser M. G., Kelsall T., Lisse C. M., Moseley S. H., Silverberg R. F., Sodroski T. J., Weiland J. L., 1995, *ApJ*, 445, 716
- Feast M., Whitelock P., 1997, *MNRAS*, 291, 683
- Flynn C., Holmberg J., Portinari L., Fuchs B., Jahreiß H., 2006, *MNRAS*, 372, 1149
- Ghez A. M., Salim S., Weinberg N. N., Lu J. R., Do T., Dunn J. K., Matthews K., Morris M. R., Yelda S., Becklin

E. E., Kremenek T., Milosavljevic M., Naiman J., 2008, *ApJ*, 689, 1044

Gillessen S., Eisenhauer F., Trippe S., Alexander T., Genzel R., Martins F., Ott T., 2009, *ApJ*, 692, 1075

Gilmore G., Reid N., 1983, *MNRAS*, 202, 1025

Gunn J. E., Gott III J. R., 1972, *ApJ*, 176, 1

Guo Q., White S., Li C., Boylan-Kolchin M., 2010, *MNRAS*, 404, 1111

Jurić M., Ivezić Ž., Brooks A., et al., 2008, *ApJ*, 673, 864

Kent S. M., Dame T. M., Fazio G., 1991, *ApJ*, 378, 131

Kerr F. J., Bowers P. F., Jackson P. D., Kerr M., 1986, *A&AS*, 66, 373

Klypin A., Zhao H., Somerville R. S., 2002, *ApJ*, 573, 597

Knapp G. R., Stark A. A., Wilson R. W., 1985, *AJ*, 90, 254

Kuijken K., Gilmore G., 1989, *MNRAS*, 239, 605

—, 1991, *ApJL*, 367, L9

Launhardt R., Zylka R., Mezger P. G., 2002, *A&A*, 384, 112

Law D. R., Majewski S. R., Johnston K. V., 2009, *ApJL*, 703, L67

Li C., White S. D. M., 2009, *MNRAS*, 398, 2177

Li Y., White S. D. M., 2008, *MNRAS*, 384, 1459

López-Corredoira M., Cabrera-Lavers A., Garzón F., Hammersley P. L., 2002, *A&A*, 394, 883

Malhotra S., 1994, *ApJ*, 433, 687

—, 1995, *ApJ*, 448, 138

McGaugh S. S., Schombert J. M., de Blok W. J. G., Zargursky M. J., 2010, *ApJL*, 708, L14

McMillan P. J., Binney J. J., 2010, *MNRAS*, 402, 934

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, *Journal of Chemical Physics*, 21, 10871092

Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563

Neto A. F., Gao L., Bett P., Cole S., Navarro J. F., Frenk C. S., White S. D. M., Springel V., Jenkins A., 2007, *MNRAS*, 381, 1450

Ojha D. K., Bienaymé O., Mohan V., Robin A. C., 1999, *A&A*, 351, 945

Picaud S., Robin A. C., 2004, *A&A*, 428, 891

Reid M. J., Brunthaler A., 2004, *ApJ*, 616, 872

Reid M. J., Menten K. M., Zheng X. W., et al., 2009, *ApJ*, 700, 137

Reylé C., Robin A. C., 2001, *A&A*, 373, 886

Robin A. C., Reylé C., Derrière S., Picaud S., 2003, *A&A*, 409, 523

Rygl K. L. J., Brunthaler A., Reid M. J., Menten K. M., van Langevelde H. J., Xu Y., 2010, *A&A*, 511, A2+

Sato M., Reid M. J., Brunthaler A., Menten K. M., 2010, *ApJ*, 720, 1055

Schmidt M., 1956, *Bull. Astr. Inst. Neth.*, 13, 15

Schönrich R., Binney J., Dehnen W., 2010, *MNRAS*, 403, 1829

Sharma S., Bland-Hawthorn J., Johnson K., J. B. J., 2011, *ApJ*, in press

Siegel M. H., Majewski S. R., Reid I. N., Thompson I. B., 2002, *ApJ*, 578, 151

Smith M. C., Ruchti G. R., Helmi A., et al., 2007, *MNRAS*, 379, 755

Spergel D. N., Malhotra S., Blitz L., 1996, in *Spiral Galaxies in the Near-IR*, D. Minniti & H.-W. Rix, ed., pp. 128–+

Springel V., White S. D. M., Jenkins A., et al., 2005, *Nat*,

435, 629

Weaver H., Williams D. R. W., 1974, *A&AS*, 17, 1Widrow L. M., Pym B., Dubinski J., 2008, *ApJ*, 679, 1239Wilkinson M. I., Evans N. W., 1999, *MNRAS*, 310, 645Xue X. X., Rix H. W., Zhao G., et al., 2008, *ApJ*, 684, 1143

## APPENDIX A: MODELS WITH DIFFERENT DISC SCALE HEIGHTS

For completeness we now provide a table of the mean and standard deviation of the parameters of our models in cases where we consider disc scale-heights which differ from our default values. These results are discussed in Section 5.1.

	$h_{d,\text{thin}}$	$h_{d,\text{thick}}$	$\Sigma_{d,0,\text{thin}}$	$R_{d,\text{thin}}$	$\Sigma_{d,0,\text{thick}}$	$R_{d,\text{thick}}$	$\rho_{b,0}$	$\rho_{h,0}$	$r_h$	$R_0$		
Mean	0.25	0.75	670	3.03	280	3.20	95.4	0.013	17.4	8.28		
Std. Dev.	-	-	123	0.24	125	0.56	7.0	0.006	3.9	0.16		
Mean	0.25	0.90	695	3.00	278	3.24	95.3	0.012	17.8	8.29		
Std. Dev.	-	-	119	0.22	120	0.53	6.9	0.005	3.9	0.16		
Mean	0.25	1.05	713	3.01	292	3.22	95.8	0.011	18.8	8.28		
Std. Dev.	-	-	124	0.23	137	0.57	6.8	0.005	5.1	0.16		
Mean	0.30	0.75	748	2.97	232	3.31	95.3	0.011	18.3	8.28		
Std. Dev.	-	-	118	0.21	113	0.60	6.8	0.005	4.2	0.16		
Mean	0.30	0.90	741	3.00	238	3.29	95.5	0.012	18.0	8.29		
Std. Dev.	-	-	123	0.22	110	0.56	6.9	0.006	4.3	0.16		
Mean	0.30	1.05	757	2.99	236	3.28	95.6	0.012	18.1	8.29		
Std. Dev.	-	-	116	0.21	103	0.55	6.8	0.005	4.0	0.16		
Mean	0.35	0.75	777	2.97	194	3.33	95.5	0.013	17.9	8.28		
Std. Dev.	-	-	127	0.21	81	0.54	6.9	0.007	4.6	0.16		
Mean	0.35	0.90	784	2.98	208	3.30	95.6	0.012	18.2	8.29		
Std. Dev.	-	-	131	0.21	87	0.55	6.9	0.006	4.4	0.16		
Mean	0.35	1.05	790	3.00	214	3.28	95.6	0.012	18.5	8.29		
Std. Dev.	-	-	122	0.21	98	0.57	6.8	0.006	5.0	0.16		
	$h_{d,\text{thin}}$	$h_{d,\text{thick}}$	$v_0$	$M_b$	$M_*$	$M_v$	$K_{z,1.1}$	$\Sigma_{d,\odot}$	$\rho_{*,\odot}$	$\rho_{h,\odot}$	$f_{d,\odot}$	$g_{*,\odot}/g_{h,\odot}$
Mean	0.25	0.75	239.0	8.96	62.7	1240	76.4	60.3	0.097	0.0107	0.140	1.24
Std. Dev	-	-	4.7	0.65	6.2	220	5.3	7.5	0.012	0.0010	0.012	0.28
Mean	0.25	0.90	239.3	8.94	64.0	1250	76.1	61.7	0.096	0.0106	0.120	1.27
Std. Dev	-	-	4.6	0.64	6.2	220	5.2	7.5	0.012	0.0010	0.012	0.28
Mean	0.25	1.05	238.9	8.99	65.7	1280	76.6	63.7	0.099	0.0104	0.101	1.36
Std. Dev	-	-	5.0	0.64	6.4	260	5.2	7.7	0.012	0.0011	0.012	0.33
Mean	0.30	0.75	238.7	8.94	63.7	1280	77.1	61.4	0.086	0.0106	0.140	1.30
Std. Dev	-	-	4.8	0.64	5.9	230	5.1	7.1	0.010	0.0010	0.012	0.29
Mean	0.30	0.90	239.2	8.96	64.3	1260	76.5	62.0	0.085	0.0106	0.120	1.29
Std. Dev.	-	-	4.8	0.65	6.3	240	5.3	7.6	0.010	0.0010	0.012	0.30
Mean	0.30	1.05	239.3	8.97	65.0	1260	76.1	62.7	0.085	0.0105	0.100	1.31
Std. Dev	-	-	4.8	0.64	6.5	220	5.2	7.7	0.010	0.0010	0.012	0.30
Mean	0.35	0.75	239.0	8.96	63.5	1260	76.5	61.1	0.077	0.0106	0.140	1.28
Std. Dev	-	-	4.9	0.65	6.3	240	5.3	7.6	0.010	0.0011	0.012	0.31
Mean	0.35	0.90	239.2	8.97	64.9	1260	76.5	62.5	0.076	0.0105	0.120	1.32
Std. Dev	-	-	4.9	0.65	6.6	240	5.4	7.8	0.010	0.0011	0.012	0.31
Mean	0.35	1.05	239.3	8.97	65.7	1270	76.4	63.5	0.077	0.0104	0.100	1.35
Std. Dev	-	-	4.8	0.64	6.8	260	5.4	8.1	0.010	0.0011	0.012	0.34

**Table A1.** Mean and standard deviation of the parameters (upper) and derived properties (lower) of our models, for various values of the thin and thick disc scale-heights ( $h_{d,\text{thin}}$  &  $h_{d,\text{thick}}$  respectively). This is very similar to Table 2. The results with our standard scale-heights ( $h_{d,\text{thin}} = 0.3$  kpc,  $h_{d,\text{thick}} = 0.9$  kpc) are included in this table as well as in Table 2. Again, distances are quoted in units of kpc, velocities in  $\text{km s}^{-1}$ , masses in  $10^9 M_\odot$ , surface densities in  $M_\odot \text{pc}^{-2}$ , densities in  $M_\odot \text{pc}^{-3}$ , and  $K_{z,1.1}$  in units of  $(2\pi G) \times M_\odot \text{pc}^{-2}$ .